Comments from anonymous Referee #2:

We would like to thank the reviewer for his/her helpful comments. We hope that we could address all questions and unclear points satisfactorily.

Legend: Author comments in blue, Referee comments in black.

Review egusphere-2023-522

Investigation of meteorological conditions and BrO during Ozone Depletion Events in Ny-Ålesund between 2010 and 2021

1 general comments

evaluating the overall quality of the discussion paper

The paper investigates the importance of polar cyclones and associated meteorological condition on ozone depleting events (ODEs) observed in the proximity of Ny-Ålesund. The authors apply complementary datasets of both ozone (O3) and BrO (e.g. in situ observations and various remote sensing products) and combine these with Lagrangian (FLEXPART) and Eulerian (WRF) modeling techniques. Though, no Chemistry Transport Model (CTM) or process modeling were applied in this study. Overall, the paper is

- well-structured,
- comprehensible in its application of the composite data analysis.

The language is

- comprehensive
- and for most parts only minor corrections may apply.

Though this paper was, overall, a real pleasure to read, a major shortcoming emerges from the distinction between background (no ODE) state of the atmosphere and ODE state. The authors define no ODE, in principle, by Y noODE = Y tot − Y gap O3 data − Y [O3]<15 ppb , where Y in this case refers to one record in 1-hourly meteorological data (e.g. Y tot would correspond to 744 records in March). Applying this definition of the background state of the atmosphere most likely reduces the explanatory power of the anomaly analysis because ODE conditions are likely to be contained in the no ODE data records. This clearly emerges from the sensitivity analysis in Section 3.2.1 where a lower threshold (20 ppb) was applied. This weakened the found dipole structure. The anomaly analysis could potentially benefit from stronger constraints on the background state, however, the general outcome of the study will probably not change significantly. Hence, either the authors, unconstrained by computational and human resources, repeat the anomaly analysis with a more clear-cut background state or discuss the implications more thoroughly in Section 3.2.1.

This topic is discussed below in p9 l219-223.

2 specific comments

individual scientific questions/issues

- p4 l116: ”[...] not necessarily capture all ODEs at sea level.”: Could you possibly give an estimate based on your composite analysis? Or give a number of coincident events?

  Due to the lack of continuous measurements in Ny-Ålesund at sea level, it is not possible to give an estimation. However, we compared the Zeppelin data with the 14 ozone sondes that measured ODEs. We found three ODEs in the ozone sonde data, which were not marked as ODE in the Zeppelin data. Two of them showed ODEs in the Zeppelin data shortly before/after the sonde launch, the third showed no sign of ODEs in the Zeppelin data at all. One out of the three ozone sondes has values above 15 ppb at the altitude of the Zeppelin station, the other two were below the 15 ppb threshold.

  When the remaining 11 ozone sondes marked as ODE were launched, ODEs were also visible in the Zeppelin data. There was no ozone sonde which was marked as no ODE, when an ODE was measured on Zeppelin mountain at the same time.

- p5 l121: ”[...] ODEs occur mainly in polar spring [...]”: Is there evidence that they do occur in, e.g., fall?

  Neither dataset shows ODEs (ozone < 15 ppb) in the fall and I am not aware of any publication on this. However, ODEs still occur from time to time in June, which can also be seen in the Zeppelin dataset (e.g., June 2013 (103 ODE hours), June 2014 (43 ODE hours)).

  To avoid confusion, the word 'mainly' has been removed from the sentence.

- p5 Table 1: Would it make more sense to differentiate between ODEs and total number of ascents in the respective month?

  This table should highlight the number of ODE and no ODE cases for different months and years, which will later form the basis of the composite analysis. Accordingly, it is more important here to show the number of ODEs and no ODEs and not to emphasize the total number of measurements. Adding additional columns with the total number of measurements would make the table too cluttered.

- p6 Table 2 (and respective paragraphs in the text): How is ”ODE” count exactly defined here? It seems as if you count consecutive hours of below threshold O3 as one event each. An ODE can, in fact, as your correctly wrote and showed (Fig 7), last for several days. Hence, your definition of ODE in this Table is not consistent with, e.g., Section 3.3.

  It is correct that consecutive hours below the threshold are each marked with the abbreviation ODE, which is not consistent with the definition of ODE. However, to avoid introducing a new abbreviation and jumping back and forth between the new abbreviation and ODE, it has been decided to still use ODE, but the following sentences have been included in Section 2.1 to avoid confusion:

  We here use the name ODE although it is not quite correct in this context, since during a longer ODE, all consecutive hours below the threshold are individually marked with the abbreviation ODE, which is not consistent with the definition of ODE. However, in order to avoid introducing a new abbreviation, we kept the term ODE when referring to individual hours having ozone values below the threshold.

In any case, no ODE and ODE [hours with O3 below threshold] do not add up to the total number of hours in each month, e.g., 744 h in March. It is not clear if this is caused by gaps in data. Could you clarify this?

These are gaps in the Zeppelin data. E.g., in March and April 2016, and there are several smaller data gaps (hours to days) spread over the years.

In order to make this clear, the following sentence has been added: *Note that Zeppelin ozone data has several data gaps that result in an incomplete number of hours of data per month.*

- p5 l122: How did you define the 15 ppb threshold? Is it possible to clearly separate populations in a histogram of O3 monitoring (bi-modal distribution)?

  Discussed below (p13 Section 3.2.1).

- p5 l129: How like do you detect ”fake” ODEs due to low hemispheric O3 background especially at higher thresholds?

  We cannot eliminate the possibility that some of the ODEs are fraudulent. Nevertheless, we think it is not likely that they were created by any mechanism other than depletion by halogens around the 15 ppb and 20 ppb threshold.

- p7 l154–177: ”From 2010 until 2017 [...]”: Too specific and technical. You may summarize the technical information about the different satellite products in a table and keep only additional information in the text.

  The entire paragraph has been removed and replaced by a table. Some information from the deleted part has been added to the paragraph above.

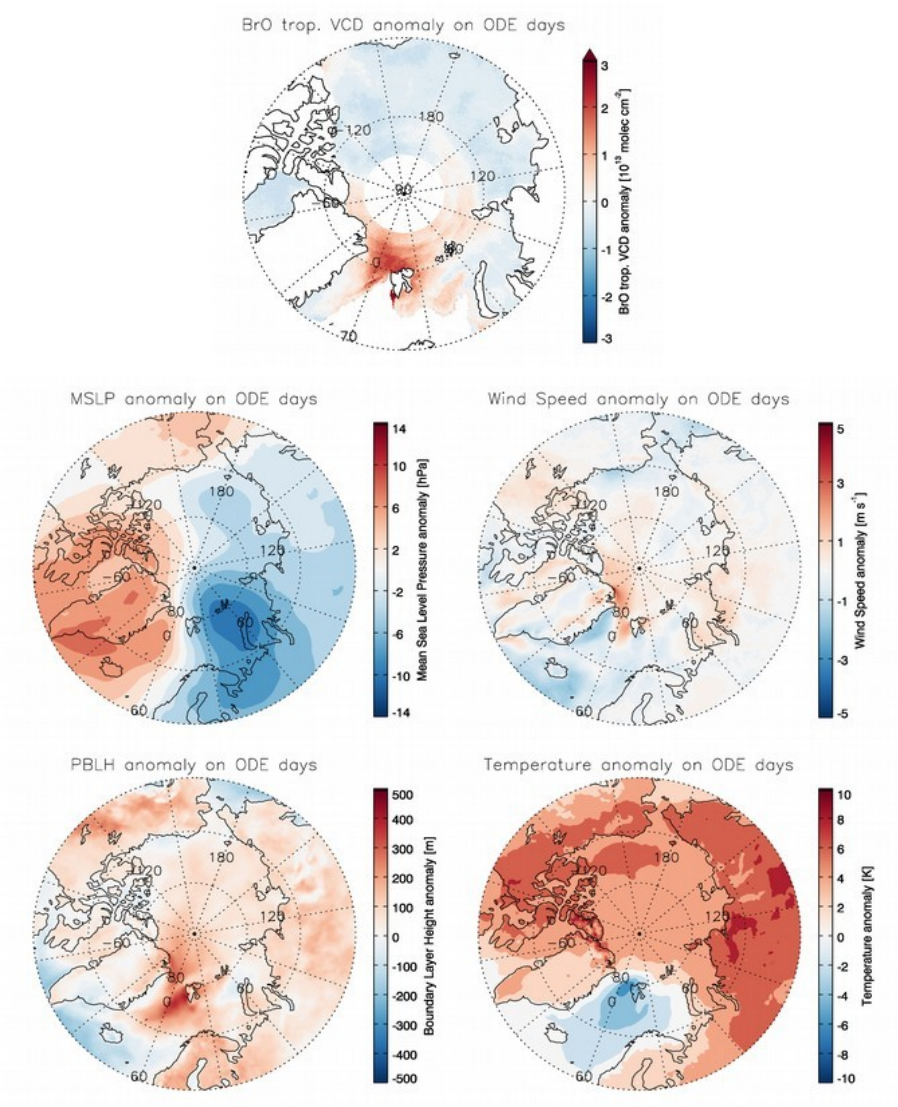
- p8 l195–196: ”[...] two domains were used in a two-way nested run, i.e. the values of the coarse domain are overwritten by values of the higher resolution [...]”: That’s not entirely correct, as changes in the inner and outer domain influence field information of one another in a two-way coupled model system. Either explain the coupling in detail or drop the second half of the sentence starting with ”i.e.”.

  The second half of the sentence has been removed.

- p9 l219-223; eq (1): Y ODEanom = Y ODE − Y noODE : This method is slightly problematic in the sense that you may have erroneously detected ODEs / noODEs in both data sets (even more so in the data sets with the lowered threshold of 20 ppb). These will blur the signal you are after. In this context, the definition of ODE that you apply to the ozone monitoring data is not sound (see also comment regarding Table 2). Consecutive hours of below threshold ozone concentrations are a necessity to identify an ”event”, hence the number of hours of ODE is misleading with respect to ODE statistics. The lag analysis that you apply in the following should be associated to the onset/end of an individual ODE. In summary, if you use times of ”no ODE” you have to make sure that these really do not contain any ODE which I currently cannot see satisfied in your analysis. Another threshold, e.g. for normal ozone concentrations, could do the trick.

  The issue of the naming ‘ODE’ has been discussed above in p6 Table 2.

  Regarding the second issues arising from the no ODE dataset erroneously containing ODEs: we performed a second run on the Zeppelin dataset, using a 15 ppb threshold for ODE (as in the article) and a 40 ppb threshold for no ODE. Any measurements between 15 and 40 ppb are not incorporated. The results are shown in the figures below:



Compared to the results shown in the paper (for the 15 ppb threshold), we can see that the overall pattern stays the same, but there are enhanced anomalies in BrO, MSLP, PBLH and temperature. Setting a no ODE threshold removes weaker ODEs, resulting in increased anomalies, but the overall pattern remains. The effect is similar to the result when changing the ODE threshold is changed.

These findings have been included into Section 3.2.1.

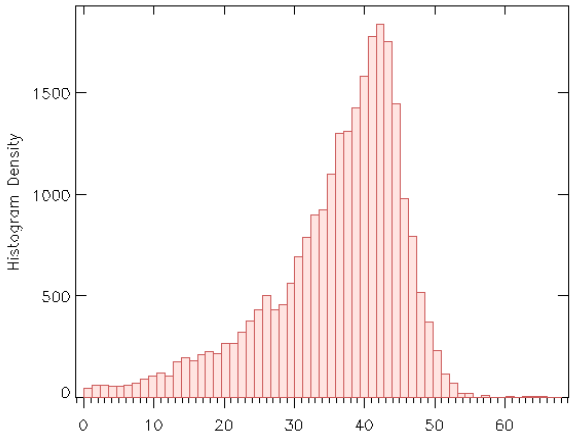
- p10 Fig 2: Regarding the shown data, where does the ”white” area south of Svalbard come from? Did you use some kind of ”sea ice edge” filtering to exclude these data or are there no data due to retrieval constraints?

  The satellite data has been sea-ice flagged to analyse only sea ice covered areas (described in Section 2.3). Hence no satellite data south of Svalbard was included in the analysis.

- p11 l274: ”long computing times”: On which kind of system? Personal computer, HPC? If former is the case an application of other computational resources should have been considered for this analysis, perhaps?

  A personal computer has been used and we agree that computing times are not an issue when using larger computing infrastructure.

- p13 Section 3.2.1: How did you derive the different threshold values? Are they ”random” choices or were the derived from a bi-modal analysis of ozone concentration distributions if that is even possible? As you show in this section, the choice of threshold is crucial for identifying a causal signal.



- As shown in the Histogram above, it is not possible to derive a clear ozone threshold from a bi-modal analysis. The thresholds are more of a ‘random’ choice, based on several applied threshold values from other studies (see Halfacre et al., 2014: Section 2.2 and supplementary). Another criterion was to get a sufficiently large number of ODEs in the ozone sonde data set.

  Halfacre, J. W., Knepp, T. N., Shepson, P. B., Thompson, C. R., Pratt, K. A., Li, B., Peterson, P. K., Walsh, S. J., Simpson, W. R., Matrai, P. A., Bottenheim, J. W., Netcheva, S., Perovich, D. K., and Richter, A.: Temporal and spatial characteristics of ozone depletion events from measurements in the Arctic, Atmos. Chem. Phys., 14, 4875–4894, https://doi.org/10.5194/acp-14-4875-2014, 2014.

  Have you analyzed coincident ODEs in both sounding and monitoring data with respect to the meteorological and BrO conditions? Or are there too few coincident ODEs?

  We have not analysed coincident ODEs in both sounding and monitoring data. However, as described in p4 l116 there are 11 out of 14 coincident ODEs, and only one ODE in sounding data does not show any signs of ODE in the Zeppelin data. Therefore, the BrO and meteorological conditions for the coincident ODEs are most likely very similar to the results shown in Figure 2 and 3 for the ozone sondes.

- p14 l309: ”Overall, the anomalies are slightly less pronounced when using the 20 ppb threshold.”: This doesn’t really come as surprise. By lowering the threshold you’d allow for more ”false positive” ODEs that can originate from both transport and mixing of air parcels with different trace gas concentrations, as well as the inclusion of subsiding ODEs where the actual cause is not present in your time lag analysis.

  It is not clear what is meant by ‘false positive’ ODEs. Even though the threshold is higher, ozone levels are still 50% below the normal background value of 40 ppb. Even if the ozone is decomposed elsewhere, the ozone level upon arrival in Ny-Ålesund must remain below a certain threshold to be considered as an ODE. Therefore, the conditions for ozone depletion must be similar or the same there as well.

It is correct that the lower threshold also includes more points in time of subsiding ODEs, leading to a slight bias. But as already mentioned above, 20 ppb is well below a normal background value, so it can be assume that the ODE conditions are still present, albeit in a subsiding form.

Even though it did not come as a surprise that the anomalies are slightly less pronounced, we have not been able to come up with a satisfactory explanation yet.

- p17 l361–370: ”Several years [...] show strong anomalies [...] other years still show similar patterns [...]”: Let’s turn this around: If you’d find similar patterns in your meteorological fields but no ODE, wouldn’t that mean that these meteorological conditions do not suffice as cause of ODEs?

  This is a legitimate question, but it is beyond the scope of this paper to answer adequately. For that, we would need to establish a routine which makes it possible to identify ODE meteorological conditions for the Arctic region and compare it to the Zeppelin ozone data. If similar meteorological conditions are recognized more frequently in spring, which however do not lead to an ODE, the influence of meteorology on ODEs would have to be reviewed again. However, it would have to be taken into account that other factors (sea ice coverage, amount of saline aerosols, etc.) can also play a role in ODEs.

- p17 l385–387: ”[...] the amount of BrO is not fully captured by the satellite observations.” Could this be due to the algorithm used to separate tropospheric and stratospheric columns?

  The stratospheric correction is probably an issue alongside the satellite's problem of detecting local phenomena due to lack of sensitivity.

  Following sentence has been changed: T*his leads to the assumption that the amount of BrO is not fully captured by the satellite observations, due to lack of sensitivity in detecting local phenomena.*

- p21 l452: ”Due to a decrease of sea ice extend less source area for BEE will be available.” Without taking the processes associated with BEE and ODE into account, this remains highly speculative. Given blowing snow on first year sea ice and brine on (young) sea ice is among the major sources of Br in the polar spring boundary layer (see e.g. 10.5194/acp-12-6237-2012 for a review of processes), the extent of the sea ice is probably less important compared to the structure and dynamics of the sea ice, higher wind speeds, and a change in frequency in the occurrence of polar lows. Dynamics of sea ice formation have been notoriously hard to detect with passive sounding satellite-born instruments, but advances might have been made in recent years.

  This sentence has been removed as it is too speculative.

3 technical corrections

purely technical corrections

If no further comment has been written, it should be considered as ‘done’.

- p2 l24: ”during sunlight”: term?

  Changed into ‘*when sunlight is present*’

- p2 R1–R6: Typesetting of chemical formulas: $Br_2 \rightarrow Br_2$, asf.

- p6 Fig 1: (left) Maybe indicate altitude of Zeppelin observatory?
- p7 l154 ff: ”From 2010 until 2017 [...]”: Duplicate of paragraph ”To analyse [...]” (p7 l146 ff). Please condense the two paragraphs.

  See above p7 l154–177
- p8 l183–185: ”Additionally, to analyse [...]”: This sentence might be grammatically incorrect. Maybe better: *Daily AMSR (...) sea ice concentration (SIC) observations on a 25 × 25 km 2 grid have been used to analyse the SIC [...].*
- p8 l184, l191, l190: Style! ”was used” is used in each of these consecutive sentences. You may want to rephrase.
- p8 l196–198 and Fig 9: Definition of WRF domains: Which projections were used? Does the clipping shown in Fig 9 represent the boundaries of the outer domain? Would it be possible to indicate the location of the inner domain in the WRF related plots in Fig 9?
- p9 l213, l214, l225, and others: ”time points” Incorrect term. Rephrase → *times*.
- p9 l215–216: ”To separate the ozone data [...]”: Repetition of Section 2.1. Remove or rephrase: *From the ozone data (see Section 2.1), we found 14 ODEs and 228 no ODEs in ozone sonde data records and [...].*
- p9 l219 ff: Typesetting of formulas. If not defined otherwise in the journal’s style guide, non-indexing subscripts should not be set in italic font: $Y_{ODE} \rightarrow Y_{\mathrm{ODE}}$.
- p10 Fig 2: ”BrO VCD anomalies for ODE and no ODE” Caption text confusing? I would assume, you should drop ”for ODE and no ODE” here. Regarding the shown data, were does the ”white” area south of Svalbard come from? Did you use some kind of ”sea ice edge” filtering to exclude these data or are there no data?

  ”for ODE and no ODE” has been removed in the caption of Figure 2 and 3.

  The second part of this comment has been answered above in p10 Fig 2.
- p10 l235: Typo: ”Ny-Alesund” → Ny-Ålesund
- p10 l242: Typo: ”the the” remove one ”the”; Missing comma after adverb?: ”Normally the Icelandic low [...]”
- p10 l243: ”Due to the lower pressure [...]”: Perhaps low pressure system? (But I’m not firm in weather synoptics.)
- p15 l314–345: ”[...] a more pronounced lower pressure anomalies [...]”: Mismatching singular article and plural noun.
- p17 l378: ”As shown in the blue line”: preposition?: *As shown by the blue line*
- p17 l379: ”in the same altitude”: preposition?: *at the same altitude*
- p21 l452: ”extend” → *extent*
- p27: Empty → Fig C1 should have appeared here.
- p25: Fig A5 is not referenced in the text.